DOCUMENT RESUME

ED 148 901 TM 006 889

TITLE	Validation of the Proficiency Examination for Diagnostic Radiologic Technology. Final Report.
INSTITUTION	Educational Testing Service, Princeton, N.J.
SPONS AGENCY	Health Resources Administration (DHEW/PHS), Bethesda, Md. Div. of Associated Health Professions.
PUB DATE	20 May 77
CONTRACT	231-75-0207; NIH-72-4226
NOTE	114p.; Not available in hard copy due to marginal reproducibility of original
EDRS PRICE	MF-$0.83 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS	Correlation; Credentials; *Cutting Scores; *Equivalency Tests; Evaluation Methods; Informal Assessment; Job Skills; Multiple Choice Tests; *Performance Tests; *Personnel Evaluation; *Radiologic Technologists; Radiology; Scoring Formulas; Skill Analysis; Supervisors; *Test Validity
IDENTIFIERS	Content Validity; Criterion Related Validity; *Proficiency Exam Diagnostic Radiologic Technology

ABSTRACT

The validity of the Proficiency Examination for Diagnostic Radiologic Technology was investigated, using 140 radiologic technologists who took both the written Proficiency Examination and a performance test. As an additional criterion measure of job proficiency, supervisors' assessments were obtained for 128 of the technologists. The resulting data indicated that the total score on the Proficiency Examination was a valid indicator of overall job proficiency for radiologic technologists, and that separate subscores do not provide additional valid information about proficiency in the specific skill areas of radiation protection, patient care, technique and quality control, and positioning. The effects of using a pass/fail cutoff at several selected levels were also investigated. The results indicated that any cutoff between 45 and 75 percent of the total possible score on the Proficiency Examination will produce systematic differences in actual performance between the groups passing and failing the Proficiency Examination. (Author/BW)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED148901

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

TM

VALIDATION OF THE PROFICIENCY EXAMINATION FOR DIAGNOSTIC RADIOLOGIC TECHNOLOGY

Final Report

Division of Associated Health Professions

Contract No. 231-75-0207

May 20, 1977

EDUCATIONAL TESTING SERVICE
PRINCETON, NEW JERSEY

TM006 889

ERRATA

The last sentence on page 49 should read as follows:

For example, the block of four numbers in the upper left corner of the table indicates that with a pass/fail cutoff of 85 percent on the Proficiency Examination, 3 technologists who were rated "nearly perfect" in all four skill areas would pass the Proficiency Examination, while 9 would fail; 14 technologists who were rated less than "nearly perfect" in some area would pass the Proficiency Examination, while 114 would fail.

A corrected Table 25 which appears on the back of this page replaces Table 25 on page 51.

Table 25

Two-way pass/fail classification tables of Proficiency Examination and supervisor's assessment based on overall judgments.

Proficiency Examination		Supervisor's Assessment: "Nearly perfect" in all areas		"Consistently good" or better in all areas		"Generally good" or better in all areas	
		Yes	No	Yes	No	Yes	No
85%	P	3	14	9	8	16	1
	F	9	114	40	83	83	40
75%	P	6	40	22	24	41	5
	F	6	88	27	67	58	36
65%	P	9	70	35	44	65	14
	F	3	58	14	47	34	27
55%	P	12	99	44	67	87	24
	F	0	29	5	24	12	17
45%	P	12	118	47	83	94	36
	F	0	10	2	8	5	5

JUL 0 1 1977

FINAL REPORT

VALIDATION OF THE PROFICIENCY EXAMINATION FOR

DIAGNOSTIC RADIOLOGIC TECHNOLOGY

May 20, 1977

Prepared By

Educational Testing Service

Princeton, New Jersey 08540

for

Division of Associated Health Professions

Department of Health, Education, and Welfare

ACKNOWLEDGEMENT

We would like to express our appreciation to all who contributed their time, effort, expertise, and enthusiasm to make this project possible. In addition to the individuals listed on succeeding pages, many others contributed significantly by serving as examinees and patient-actors, by assisting in the preparation of printed materials, and by providing helpful comments, criticisms, and suggestions.

E.T.S. STAFF

James E. Alwood, M.S.
Senior Staff Systems Analyst

William Kastrinos, Ph.D.
Senior Examiner

Samuel A. Livingston, Ph.D.
Associate Program Research Scientist

Amiel T. Sharon, Ph.D.
Program Director

H.E.W. PROJECT OFFICER

Margaret A. Wilson, Ph.D.
Educational Specialist
Division of Associated Health Professions
Bureau of Health Manpower

ADVISORY COMMITTEE

Michael J. Audet, R.T.
Bureau of Radiological Health

Loy T. Brown, M.D.
American College of Radiology

Ralph W. Coates, R.T.
American Registry of
Radiologic Technologists

Marilyn Fay, R.T.
American Society of
Radiologic Technologists

Charles A. Jacobi, R.T.
American Registry of Clinical
Radiography Technologists

Carole Palmer, R.T.
American Society of
Radiologic Technologists

PERFORMANCE TEST DEVELOPMENT TEAM

John Dwyer, M.D.
Helene Fuld Medical Center
Trenton, New Jersey

Mary Vahey, R.T.
Helene Fuld Medical Center
Trenton, New Jersey

Samuel A. Livingston, Ph.D.
Educational Testing Service
Princeton, New Jersey

SUPERVISOR'S ASSESSMENT DEVELOPMENT TEAM

Richard L. Barach, M.D.
Medical Center
Princeton, New Jersey

William J. Snyder, R.T.
Medical Center
Princeton, New Jersey

Samuel A. Livingston, Ph.D.
Educational Testing Service
Princeton, New Jersey

PERFORMANCE TEST SCORING COMMITTEE MEMBERS

Committee I:

Mary Vahey, R.T., Chairperson
Helene Fuld Medical Center
Trenton, New Jersey

Georgene B. McKenzie, R.T.
Jersey City Medical Center
Jersey City, New Jersey

Patricia A. Myers, R.T.
Montefiore Hospital
Pittsburgh, Pennsylvania

Committee II:

Mary Vahey, R.T., Chairperson
Helene Fuld Medical Center
Trenton, New Jersey

Georgene B. McKenzie, R.T.
Jersey City Medical Center
Jersey City, New Jersey

Nancy Pae, R.T.
Helene Fuld Medical Center
Trenton, New Jersey

Charlotte Todesco, R.T.
Franklin Square Hospital
Baltimore, Maryland

PERFORMANCE TEST ADMINISTRATORS

George P. Osmun, R.T.
Mt. Sinai Hospital
Hartford, Connecticut

Clemens P. Pearson, R.T.
St. Joseph's Hospital
Phoenix, Arizona

Susan M. Foster, R.T.
Medical College of Georgia
Augusta, Georgia

Lester E. Lepley, R.T.
Franklin Square Hospital
Baltimore, Maryland

Sr. Mary Rose Mitchell, R.T.
Essex County College
Newark, New Jersey

Georgene B. McKenzie, R.T.
Jersey City Medical Center
Jersey City, New Jersey

Albert M. Snopek, R.T.
Middlesex County College
Edison, New Jersey

Diane M. Neiner, R.T.
St. Vincent Health Center
Erie, Pennsylvania

Patricia A. Myers, R.T.
Montefiore Hospital
Pittsburgh, Pennsylvania

Frank E. Ellis, R.T.
Williamsport Hospital
Williamsport, Pennsylvania

Lawrence B. Kirschner, R.T.
Medical College of Virginia
Richmond, Virginia

Neta B. McKnight, R.T.
Jackson State Community College
Jackson, Tennessee

Riley F. Trimm, R.T.
Veterans Administration Hospital
Birmingham, Alabama

Nicholas R. Barraco, R.T.
Veterans Administration School
of Radiologic Technology
Westhaven, Connecticut

Marilyn D. Holland, R.T.
Mercy Hospital
Iowa City, Iowa

Sandra E. Harrison, R.T.
Pitt County Memorial Hospital
Greenville, North Carolina

PERFORMANCE TEST OBSERVERS

George P. Osmun, R.T.
Mt. Sinai Hospital
Hartford, Connecticut

Nancy A. Gurry, R.T.
Mt. Sinai Hospital
Hartford, Connecticut

Clemens P. Pearson, R.T.
St. Joseph's Hospital
Phoenix, Arizona

Michele Vaudreuil, R.T.
St. Joseph's Hospital
Phoenix, Arizona

J. C. Goodner, R.T.
Veterans Administration Hospital
Birmingham, Alabama

R. Gregory Carroll, R.T.
Veterans Administration Hospital
Birmingham, Alabama

Nicholas R. Barraco, R.T.
Veterans Administration School
of Radiologic Technology
Westhaven, Connecticut

John Santarsiero, R.T.
Veterans Administration School
of Radiologic Technology
Westhaven, Connecticut

Susan M. Foster, R.T.
Medical College of Georgia
Augusta, Georgia

Nancy L. Lavin, R.T.
Medical College of Georgia
Augusta, Georgia

Marilyn D. Holland, R.T.
Mercy Hospital
Iowa City, Iowa

Jarold Komro, R.T.
Mercy Hospital
Iowa City, Iowa

Lester E. Lepley, R.T.
Franklin Square Hospital
Baltimore, Maryland

Sr. Mary Rose Mitchell, R.T.
Essex County College
Newark, New Jersey

Norman Bossie, R.T.
Essex County College
Newark, New Jersey

Georgene B. McKenzie, R.T.
Jersey City Medical Center
Jersey City, New Jersey

Elizabeth Curtis, R.T.
Jersey City Medical Center
Jersey City, New Jersey

Albert M. Snopek, R.T.
Middlesex County College
Edison, New Jersey

Robert J. Chamberlain, R.T.
Middlesex County College
Edison, New Jersey

Patricia A. Copenhaven, R.T.
St. Vincent Health Center
Erie, Pennsylvania

Diane M. Neiner, R.T.
St. Vincent Health Center
Erie, Pennsylvania

Patricia A. Myers, R.T.
Montefiore Hospital
Pittsburgh, Pennsylvania

Paula M. Hermer, R.T.
Montefiore Hospital
Pittsburgh, Pennsylvania

Frank E. Ellis, R.T.
Williamsport Hospital
Williamsport, Pennsylvania

Thomas C. Schnars, R.T.
Williamsport Hospital
Williamsport, Pennsylvania

Lawrence B. Kirschner, R.T.
Medical College of Virginia
Richmond, Virginia

Andre D. Rhodes, R.T.
Medical College of Virginia
Richmond, Virginia

Jesse Jones, R.T.
Jackson State Community College
Jackson, Tennessee

Virginia Wilson, R.T.
Jackson State Community College
Jackson, Tennessee

Neta B. McKnight, R.T.
Jackson State Community College
Jackson, Tennessee

Patsy Davidson, R.T.
Jackson State Community College
Jackson, Tennessee

Elmer Hunt, R.T.
Jackson State Community College
Jackson, Tennessee

Charlotte Butler, R.T.
Jackson State Community College
Jackson, Tennessee

Sandra E. Harrison, R.T.
Pitt County Memorial Hospital
Greenville, North Carolina

Linda N. Worthington, R.T.
Pitt County Memorial Hospital
Greenville, North Carolina

Judy L. Rivenbark, R.T.
Pitt County Memorial Hospital
Greenville, North Carolina

Lynette H. Woolard, R.T.
Pitt County Memorial Hospital
Greenville, North Carolina

Marcia G. Leggett, R.T.
Pitt County Memorial Hospital
Greenville, North Carolina

Garrie W. Moore, R.T.
Pitt County Memorial Hospital
Greenville, North Carolina

CONTENTS

	Page
Introduction	1
Method	4
Criterion Measures	4
Performance test	4
Supervisors' assessments	6
Participants	9
Test Administration and Scoring	12
Results and Discussion	17
Evidence of Validity	17
Composite criterion measure	18
Separate criterion measures	19
Validity for credentialed and non-credentialed technologists	24
Results of an alternate scoring formula for the performance test	31
Usefulness of Separate Skill Area Scores on the Proficiency Examination	33
Information Concerning Pass/Fail Cutoff Scores	37
Summary and Conclusions	52
Appendices	53

Tables and Figures

Page

Tables

Table 1. Number of items or checkpoints contributing to each skill area 9

Table 2. Number of technologists in the sample who do not perform general diagnostic exams in each of five anatomical areas 11

Table 3. Number of non-credentialed vs. credentialed technologists who took the performance test at vs. away from their place of employment 13

Table 4. Number of checkpoints receiving each scoring weight in each skill area 14

Table 5. Inter-observer agreement of performance test checkpoints 15

Table 6. Correlations of Proficiency Examination with equally-weighted composite criterion, for those technologists for whom supervisor's assessment data were available . . 19

Table 7. Correlations of Proficiency Examination with performance test 20

Table 8. Correlations of Proficiency Examination with supervisor's assessments, for those technologists for whom assessment data were available 21

Table 9. Intercorrelation of criterion measures, for all technologists for whom supervisor's assessment data were available 23

Table 10. Correlations of Proficiency Examination with performance test, for credentialed and non-credentialed technologists tested at and away from their workplaces 24

Table 11. Means and standard deviations on performance test and on Proficiency Examination, for credentialed and non-credentialed technologists taking the performance test at and away from their place of employment 26

Table 12. Correlations of Proficiency Examination with supervisor's assessment based on specific situations, for non-credentialed and credentialed technologists separately 27

Tables and Figures
(continued)

Tables		Page
Table 13.	Correlations of Proficiency Examination with supervisor's assessment based on a single overall judgment for each skill area, for non-credentialed and credentialed technologists separately	28
Table 14.	Means and standard deviations on supervisor's assessments, for non-credentialed and credentialed technologists for whom supervisor's assessment data were available	30
Table 15.	Correlations of Proficiency Examination with performance test, using alternate scoring system for performance test	31
Table 16.	Correlations of performance test with supervisor's assessments, using alternate scoring system for performance test, for all technologists for whom supervisor's assessment data were available	32
Table 17.	Correlations of total score and subscores on Proficiency Examination with total score and subscores on performance test, for all participating technologists	34
Table 18.	Correlations of total score and subscores on Proficiency Examination with total score and subscores on supervisor's assessment based on specific situations, for all technologists for whom supervisor's assessment data were available	35
Table 19.	Correlations of total score and subscores on Proficiency Examination with total score and individual items of supervisor's assessment based on overall judgments, for all technologists for whom supervisor's assessment data were available	36
Table 20.	Effects of selected pass/fail cutoff scores: Proficiency Examination total score	38
Table 21.	Two-way distribution of scores on Proficiency Examination and performance test: all participating technologists	40
Table 22.	Two-way pass/fail classifications of Proficiency Examination and performance test: all participating technologists	42
Table 23.	Two-way pass/fail classification of Proficiency Examination and performance test, using alternate scoring system for performance test	44

Page

Tables

Table 24. Summary of supervisor's assessments based on overall judgments: all technologists for whom supervisor's assessment data were available 50

Table 25. Two-way pass/fail classification tables of Proficiency Examination and supervisor's assessment based on overall judgments 51

Table F-1. Intercorrelations and alpha reliabilities of Proficiency Examination scores, for all participating technologists 70

Table F-2. Effects of selected pass/fail cutoff scores: Proficiency Examination subtest score for radiation protection 71

Table F-3. Effects of selected pass/fail cutoff scores: Proficiency Examination subtest score for patient care 72

Table F-4. Effects of selected pass/fail cutoff scores: Proficiency Examination subtest score for technique and quality control 73

Table F-5. Effects of selected pass/fail cutoff scores: Proficiency Examination subtest score for positioning 74

Table F-6. Two-way pass/fail classifications of Proficiency Examination and performance test: non-credentialed technologists only 75

Table F-7. Two-way pass/fail classification of Proficiency Examination and performance test: credentialed technologists only 76

Table F-8. Two-way pass/fail classification of Proficiency Examination and performance test, using alternate scoring system for performance test: non-credentialed technologists only 77

Table F-9. Two-way pass/fail classification of Proficiency Examination and performance test, using alternate scoring system for performance test: credentialed technologists only 78

Tables and Figures
(continued)

Page

Figures

Figure 1. Pass rates for high-performance and low-performance groups: performance test cutoff = 75% 46

Figure 2. Pass rates for high-performance and low-performance groups: performance test cutoff = 65% 46

Figure 3. Pass rates for high-performance and low-performance groups: performance test cutoff = 55% 47

Figure 4. Pass rates for high-performance and low-performance groups: performance test cutoff = 45% 47

Appendices

		Page
Appendix A.	Content specifications for the Proficiency Examination	53
Appendix B.	Letter from ETS to radiology departments, and accompanying materials	56
Appendix C.	List of Workshop Participants and schedule of activities	63
Appendix D.	Missing data estimation procedure	66
Appendix E.	Scoring procedure for the performance test	68
Appendix F.	Supplementary tables	70

INTRODUCTION

This report describes and discusses the method and the findings of a research study to assess the concurrent validity of the proficiency examination for diagnostic radiologic technology, produced by Educational Testing Service (ETS) in 1972-74 under contract No. NIH 72-4226 with the Division of Associated Health Professions, U. S. Department of Health, Education and Welfare. Four forms of the proficiency examination were developed, each consisting of 230 multiple-choice questions. The four forms were constructed to be equivalent in content and difficulty. Although only one of the four forms (Form VDR2) was administered in the present study, the similarity of the four forms should justify the generalization of most of the results of this study to the other three forms as well.

Psychologists generally recognize three basic types of validity: content validity, criterion-related validity, and construct validity.[1] Two of these, content validity and criterion-related validity, are especially relevant to the proficiency examination for diagnostic radiologic technologists. Content validity in this case concerns inferences from the sample of knowledge tested by the examinations to the larger domain of knowledge necessary to perform the work of a diagnostic radiologic technologist. Criterion-related validity concerns inferences from the behavior of answering multiple-choice questions to the behavior of operating X-ray equipment, assisting and positioning patients, and other behaviors which make up the actual work of the technologist.

---

[1] For a discussion of the three types of validity, see the American Psychological Association's Standards for Educational and Psychological Tests. Washington, D.C.: American Psychological Association, 1974, pp.25-31.

The content validity of the proficiency examinations is implied by the manner in which they were constructed. A committee of experts in the field conducted a thorough review and consolidation of several task analyses,[2] which served as the basis for content specifications for the examination. Multiple-choice questions were then written and assembled into test forms according to these specifications. However, the criterion-related validity of the examinations still remained to be demonstrated. A valid proficiency examination should yield high scores for proficient technologists, including those who may be weak in academic ability or in theoretical knowledge not actually necessary to perform the job properly. It should result in lower scores for less proficient technologists, including those who may have considerable "book knowledge" and "test wiseness". The purpose of this criterion-related validity study was to provide evidence concerning the usefulness of the proficiency examinations as an indicator of proficient on-the-job performance.

Within the category of criterion-related validity, psychologists commonly distinguish between concurrent validity and predictive validity. Concurrent validity concerns inferences from the examination to actual proficiency at approximately the same time; predictive validity concerns inferences from the examination to actual proficiency at some future time. The intended use of the proficiency examinations is for the assessment of the proficiency of diagnostic radiologic technologists at the time of testing, rather than at some later time. Therefore, concurrent validity, rather than predictive validity, is of primary importance for these examinations.

---

[2] Delineation of Roles and Functions of Diagnostic Radiologic Technology Personnel and Development of Proficiency Tests. Final Report, Bureau of Health Resources Development, Contract No. NIH 72-4226. Princeton, New Jersey: Educational Testing Service, 1974.

The content specifications for the Proficiency Examination were based on a categorization of the knowledge and skills of the profession into seven major areas:[3]

1. Radiation protection
2. Technique and quality control
3. Patient care
4. Positioning
5. Ethics
6. Radiographic equipment and operator maintenance
7. Special techniques

The first four of these content areas were to account of 82% of the content of the Proficiency Examination. In order to provide more information about the validity of the Proficiency Examination, the criteria were constructed to yield separate "skill area scores" corresponding to the first four content areas listed above. These skill area scores were to serve as criteria for the portions of the Proficiency Examination that tested the corresponding content areas.

---

[3]Appendix A, page 53, presents these content specifications in greater detail.

METHOD

The first step in the validation process was to select an advisory committee of six experts in the fields of radiology and radiologic technology. The committee consisted of persons nominated by five professional organizations in these fields. One committee member was a radiologist; the other five were radiologic technologists. The organizations were asked to nominate persons who had not participated in the development of the Proficiency Examination, and none of the committee members had done so. (In fact, none had seen the Proficiency Examination in any form.)

Criterion Measures

Performance test. The next step in the validation process was to develop a "hands-on" performance test, to provide a direct measure of proficiency with which to compare the results of the Proficiency Examination. The Advisory Committee originally selected eight radiographic procedures to provide a sample of the skills needed by a Level III radiologic technologist.[4] This list of eight procedures was later shortened to six because of time constraints and other difficulties (e.g., disagreement between hospitals as to the preferred procedure, difficulty of training persons to take the role of the patient, etc.). The resulting list consisted of the following six radiographic procedures:

1. Intravenous pyelogram: scout radiograph, preparation of injection materials, and 5 minute radiograph.
2. Upper G.I. Series: posterior-anterior and right anterior oblique radiographs.

---

[4]See Delineation of Roles and Functions of Diagnostic Radiologic Technology Personnel and Development of Proficiency Tests for a definition of Level III radiologic technologist.

3. Maxilla: Waters, Caldwell, and lateral radiographs.
4. Chest: posterior-anterior and lateral radiographs.
5. Cervical spine: radiograph for detection of possible fracture.
6. Femoral neck: anterior-posterior and lateral radiographs in operating room, using mobile x-ray unit.

The actual construction of the performance test was accomplished by a team of three persons: a radiologist, a radiologic technologist, and an ETS testing specialist with experience in performance testing. The radiologic technologist was nominated by the president of the American Society of Radiologic Technologists (ASRT); the radiologist was chosen by the technologist.

The technique by which the performance test was constructed requires each procedure to be analyzed into many small, discrete actions. At each stage of the procedure, there may be one or more ways in which a technologist can make a mistake. Each of these possible mistakes is then expressed in the form of a statement of the correct action. These statements are called "checkpoints". The checkpoints are arranged in approximate sequence to form a checklist with which an observer can record the technologist's performance by checking spaces marked "yes" or "no" next to each statement. The number of checkpoints on each checklist was as follows:

Intravenous pyelogram	42
Upper G.I.	38
Maxilla	47
Chest	36
Cervical spine	19
Femoral neck (O.R.)	33

The advantage of this method of constructing performance tests is its objectivity; it leaves relatively little room for subjective judgment on the part of the observer. Therefore, the examinee's score does not depend heavily on the observer's tendency to be strict or lenient.

Administrative, safety, and legal considerations ruled out the use of actual patients in the performance test. For one of the problems (femoral neck) the "patient" was a "phantom" radiographic dummy. For the other five problems the "patients" were persons recruited and trained to take the role of the patient. (These persons were not radiologic technologists or radiologic technology students.) For some of the problems the sex and approximate age of the "patient" were specified. To avoid exposing the "patient" to radiation, the examinees were prevented from making any actual radiographic exposure. Therefore, the performance test was basically a test of the examinee's procedure in preparing the patient and materials, positioning the patient and the x-ray tube, and setting the machine controls. Each of the first five problems also included a section in which the examinee was shown one or more finished radiographs with the technical factors indicated and asked to specify any necessary corrections in technique or positioning. (There were 9 such pre-made radiographs.)

Before the performance test was administered, it was reviewed by the Advisory Committee.

Supervisors' assessments. Although the performance test is an appropriate criterion measure for a validation study of a written proficiency examination, it has some limitations. It includes a relatively small sample of the many tasks the technologist must master. It measures performance at a single time, rather than typical performance over a longer period. And the testing situation differs

from the day-to-day work situation in ways that may affect performance (e.g., the presence of the observers). For these reasons a second criterion measure was developed to supplement the performance test. This supplementary criterion was an assessment form to be completed by the technologist's regular work supervisor. The supervisor's assessment reflected the day-to-day performance of the technologist in his or her regular work situation. Therefore, despite the subjectivity and the lack of standardization inherent in this method of assessment, it provided a useful supplementary criterion measure.

The supervisor's assessment form was constructed by a team of three persons: a radiologist, a radiologic technologist, and an ETS testing specialist. The radiologist was nominated by the executive director of the American College of Radiology; the technologist was chosen by the radiologist.

One common problem with the use of rating forms in personnel evaluation is "halo effect" - the tendency of raters to form a general impression of each worker, disregarding the worker's specific strengths and weaknesses. The assessment form was constructed in a way that reflects a deliberate attempt to minimize "halo effect" - in particular, to encourage the supervisor to evaluate the technologist separately in the four skill areas of radiation protection, technique and quality control, patient care, and positioning. The major part of the form consists of items that ask the supervisor to judge the technologist's

ability to do highly specific tasks (e.g., "chest exmination on a two month old infant with respiratory distress"). The supervisor responds to each item by checking one of four statements:

1. Cannot perform this task.
2. Can perform this task only under direct supervision.
3. Can perform this task without direct supervision only if expert assistance is available.
4. Can perform this task without expert assistance or supervision.

Each specific task was selected to present particular difficulties in one of the four skill areas (see Table 1, page 9). There were four such items for radiation protection, seven for patient care, eight for technique and quality control, and nine for positioning. Four items at the end of the form called for overall judgments of the technologist's performance in the four skill areas, and two additional items measured other aspects of the technologist's performance.

Before the supervisor's assessment form was used to collect data for the validation study, it was reviewed by the Advisory Committee.

The radiologic technologist and the radiologist who developed the supervisor's assessment form also classified the items on the Proficiency Examination into the four skill areas. Several of the items were not included in any of the skill areas, and a few were included in more than one skill areas, but most of the items were included in one and only one skill area.

Table 1 shows the number of items contributing to each skill area score on each of the three measuring instruments.

Table 1

Number of items or checkpoints contributing to each skill area score.

Skill areas:	Measurement techniques: Proficiency Examination	Performance Test*	Supervisor's Assessment (items 1-28)
Radiation protection	23	27	4
Patient care	22	58	7
Technique and quality control	80	92	8
Positioning	47	76	9

*Does not include checkpoints deleted because of poor inter-observer agreement or because of excessive missing data.

Participants

The next step in the validation process was to recruit an appropriate group of radiologic technologists to serve as subjects for the study. Most of these technologists were recruited through the hospitals at which they were employed, by the following procedure. ETS requested and received from the American Society of Radiologic Technologists (ASRT) a list of hospital radiology departments that might be able and willing to serve as testing sites. A similar list was provided by the Bureau of Radiological Health (BRH) and another by the technologist who developed the performance test. ETS sent the departments on these lists a letter requesting their participation in the study.[5] Those departments that responded

[5]Appendix B, page 56, contains a copy of this letter and the materials that accompanied it.

favorably were chosen as testing sites. Two experienced technologists from each department were recruited to serve as observers for the performance test. The subjects for the study were recruited from the remaining technologists on the hospital staff. Participation in the study was voluntary, and each subject was paid fifty dollars for participating.

Because of the intended purpose of the Proficiency Examination, the sample of subjects was to include adequate numbers of both "credentialed" and "non-credentialed" technologists. "Credentialed" technologists were defined as those who had achieved certification by the American Registry of Radiologic Technologists (ARRT). All other technologists, even those licensed by a state, were considered "non-credentialed".[6] The recruitment procedure described above provided an adequate number of credentialed technologists, but very few non-credentialed technologists. The participation of additional non-credentialed technologists was obtained by contacting selected radiologic technologists in several geographic locations and asking their assistance in recruiting non-credentialed technologists to serve as subjects.[7] Most of these non-credentialed technologists had to be tested at hospitals other than those at which they were employed.

The resulting sample of subjects included 76 credentialed technologists and 64 non-credentialed technologists. Of the credentialed technologists, 12 percent had less than one year of experience, while 26 percent had more than five years of experience; the corresponding figures for the non-credentialed technologists were 11 percent and 45 percent. Table 2 is based on a question

---

[6]Eight of the participating technologists were state-licensed but not credentialed by ARRT.

[7]These contacts were made through the technologists who served as observers for the performance test, the technologist who developed the performance test, and the ASRT's Director of Continuing Education.

asking the technologists to indicate the approximate percent of their work week that they spent performing general diagnostic exams in five anatomical areas (each of which corresponded to one of the first five problems on the performance test). The proportion of the technologists not performing diagnostic exams in each of the five anatomical areas varied from 9 percent (chest) to 32 percent (genital-urinary tract). Only eight of the technologists spent over 90 percent of their work week performing diagnostic exams in a single anatomical area; all of these were non-credentialed technologists.

Table 2

Number of technologists in the sample who do not perform general diagnostic exams in each of five anatomical areas.

	Non-credentialed (n = 64)	Credentialed (n = 76)	Total (n = 140)
Genital-urinary tract	22 (34%)	23 (30%)	45 (32%)
Gastro-intestinal tract	20 (31%)	20 (26%)	40 (29%)
Skull and facial bones	13 (20%)	14 (18%)	27 (19%)
Chest	5 (8%)	8 (11%)	13 (9%)
Skeletal system	9 (14%)	11 (14%)	20 (14%)

Test Administration and Scoring

The performance test was administered at 16 different institutions located in 11 different states.[8] To prepare the observers and administrators for the actual performance testing, ETS conducted a one-day workshop in the use of the performance test and its purpose in this research study. The observers were given the opportunity to practice using each of the performance checklists and to raise questions about the administration of the performance test. Some minor revisions to the performance checklists were made at the conference as a result of suggestions by the workshop participants. Two of the developers of the performance test (the radiologic technologist and the ETS testing specialist) were available to answer questions, as were the HEW Project Officer and the ASRT Director of Continuing Education. The conference was held at a community college located near ETS and was organized by one of the performance test observers who was a faculty member at that college.[9]

Personnel and equipment limitations made it impossible for each participating technologist to take the performance test at the institution where he or she was employed. Therefore, 18 of the 76 credentialed technologists and 50 of the 64 non-credentialed technologists took the performance test at institutions other than those where they were employed, as shown in Table 3. Because equipment and procedures differ somewhat from one institution to another, these technologists may have been at a disadvantage on the performance test. This possibility has been taken into account in analyzing and reporting the results of the performance test.

---

[8]Alabama, Arizona, Connecticut, Georgia, Iowa, Maryland, North Carolina, New Jersey, Pennsylvania, Tennessee, and Virginia.

[9]Appendix C, page 63, contains a list of the persons attending the conference and a copy of the schedule of activities.

Table 3

Number of non-credentialed vs. credentialed technologists who took the performance test at vs. away from their place of employment.

	Non-credentialed	Credentialed	Combined
Tested at work	14	58	72
Tested away	50	18	68
Combined	64	76	140

The scoring formulas for the performance test were determined by two scoring committees. The first committee established scoring rules for the sections of the performance test that required the examinee to make technical judgments of finished radiographs. The second committee established a set of scoring weights to be used in combining all the checkpoints into a total score and four skill area scores. Both scoring committees were selected and chaired by the technologist who served on the performance test development team. The first committee consisted of the chairperson and two others; the second committee consisted of the chairperson and three others (two of whom had served as observers for the performance test).

The scoring weights determined by the second committee indicated the relative importance of each checkpoint (on a scale of 1 to 3) for each skill area. A weight of 3 was assigned to each checkpoint that indicated a procedure necessary to safeguard the well-being of the patient. A weight of 2 was

assigned to each checkpoint that indicated a procedure necessary to produce a usable radiograph or a procedure highly desirable for the well-being of the patient. A weight of 1 was assigned to each of the remaining checkpoints. Table 4 shows the number of checkpoints receiving each scoring weight in each skill area. (Checkpoints that were deleted because of poor inter-observer agreement or excessive missing data are not included in Table 4.) In addition, some checkpoints were considered critical; any examinee who failed one of these checkpoints received a score of zero for the entire problem. A critical checkpoint was defined as one on which an error would be serious enough to indicate incompetency by the radiologic technologist.

Table 4

Number of checkpoints receiving each scoring weight in each skill area.

	Scoring weight		
Skill area:	1	2	3
Radiation protection	3	2	22
Patient care	18	9	32
Technique and quality control	34	44	14
Positioning	13	50	14

Note: This table does not include those checkpoints deleted because of poor inter-observer agreement or excessive missing data.

Scoring of the performance test presented at least two problems not encountered in a written test: disagreement between observers and missing observations. These problems were handled by the following procedure. First, the proportion of agreement between observers was computed for each checkpoint. This statistic is simply the percentage of examinees for whom both observers checked the same response (yes or no), excluding any examinees for whom one or both observations were missing. Any checkpoint for which the inter-observer agreement was less than 85 percent was deleted from the performance test. There were only 20 such checkpoints out of a total of 217 checkpoints. Table 5 presents a summary of the inter-observer agreement for all checkpoints in the performance test.

Table 5

Inter-observer agreement of performance test checkpoints.

Inter-observer Agreement	Number of Checkpoints
100%	10
99%	20
98%	30
97%	19
96%	26
95%	20
94%	16
93%	9
92%	8
91%	12
90%	6
89%	1
88%	8
87%	7
86%	3
85%	2
less than 85%	20*

*These checkpoints were deleted from the performance test.

Next, the observations of the two observers were combined. Thus, if both observers checked "yes", the examinee received a score of 1 for that checkpoint; if both observers checked "no", the examinee received a score of 0; if one observer checked "yes" and the other checked "no", the examinee received a score of 1/2. If one observer left a checkpoint blank, the other observer's response determined the examinee's score for that checkpoint.

Where both observers left a checkpoint blank for an examinee, the missing piece of information was estimated by a procedure that takes into account both the overall performance of the examinee on that problem and the overall difficulty of the checkpoint.[10] This procedure was used when the missing data resulted from an oversight by the observers or from problems with the equipment. It was not used in cases where the examinee did not attempt a problem or was prevented by the observers from finishing a problem.[11] (A similar procedure was used for missing data on the supervisor's assessment form.)

Finally, skill area scores and total scores were computed. The total scores were computed by a formula that gives equal weight to each of the six problems, yielding scores on a scale of zero to six.[12] This formula was not used in computing the skill area scores.

---

[10]Appendix D, page 66, describes the missing data estimation procedure.

[11]Problem 5 involved a patient with a possible spinal injury. If the examinee attempted to remove the patient from the stretcher or moved the patient's head during positioning, the observers ended the problem at that point. There were 20 such examinees.

[12]Appendix E, page 68, presents a detailed, step-by-step description of the performance test scoring procedure.

RESULTS AND DISCUSSION

The information collected in this study provides evidence concerning three important questions about the Proficiency Examination:

1. Is the Proficiency Examination a valid measure of proficiency for diagnostic radiologic technologists?
2. Do the Proficiency Exmaination skill area scores provide useful and valid information that is not provided by the total score?
3. What are the effects of various possible pass/fail cutoff scores on the Proficiency Examination?

Evidence of Validity

The validity of the Proficiency Examination as a measuring instrument is indicated by the extent to which it agrees with other, more direct measures of proficiency. The statistic used to indicate the relationship between the various measures of proficiency in this study is the correlation coefficient.[13] This statistic tells the extent to which two measures agree, in the sense that persons who score above average on one measure also score above average on the other, expressed on a scale from +1 (perfect agreement) to -1 (perfect disagreement). It takes into account not only whether each person scores above or below average on the two measures, but also how far above or below average the person's scores are. Therefore it is a general measure of agreement over the full range of scores of the persons tested.

---

[13] More precisely, the Pearson product-moment correlation coefficient.

Composite criterion measure. Each of the criterion measures - the performance test and the two types of supervisor's assessment - provides a somewhat different type of information about the technologist's proficiency. These three measures can be combined into a single composite criterion by first standardizing the scores on each criterion measure and then summing the three standardized scores. The correlation of the Proficiency Examination score with this composite criterion then provides a measure of validity that takes into account the information provided by all three criteria.[14] This procedure was done for the total score and for each skill area, for those technologists for whom supervisor's assessment data were available. The resulting correlations are presented in Table 6. They indicate that the Proficiency Examination is a valid measure of overall proficiency and that the skill area scores are valid measures of proficiency in their respective areas.

---

[14]The resulting correlation is not a multiple correlation. A multiple correlation would have been misleadingly high, because the multiple correlation procedure forms an unequally weighted composite by combining variables in such a way as to maximize the resulting correlation.

Table 6

Correlations of Proficiency Examination with equally-weighted composite criterion, for those technologists for whom supervisor's assessment data were available (n = 128).

Total score	.43*
Radiation protection	.37*
Patient care	.37*
Technique and quality control	.45*
Positioning	.52*

*Different from zero at the .01 level of statistical significance.

Separate criterion measures. Table 7 presents the correlations of the Proficiency Examination with the performance test, for the total scores and for the four skill areas. Because the technologists who took the performance test at their regular workplace had an advantage over those who did not,[15] the correlations have been computed separately for these two groups, as well as for the combined group. These correlations indicate a definite tendency for the

[15]Some of the technologists tested away from their workplace may not have been familiar with the specific x-ray equipment and standard operating procedures of the hospitals at which they were tested.

Proficiency Examination and the performance test to agree in the group tested away from their workplaces and in the combined group. This finding implies that the Proficiency Examination is a valid measure of proficiency. The relationship between written and performance tests is weaker in the group tested at work, probably because there was less variation among the members of this group.[16] Among the skill areas, the relationship between the Proficiency Examination and the performance test is greatest for positioning, in both groups of technologists and in the combined group.

Table 7

Correlations of Proficiency Examination with performance test.

	Technologists tested at work (n = 72)	Technologists tested away (n = 68)	Combined (n = 140)
Total	.20	.48*	.42*
Radiation protection	.28	.49*	.43*
Patient care	.24	.12	.19
Technique and quality control	.07	.45*	.37*
Positioning	.40*	.65*	.61*

*Different from zero at the .01 level of statistical significance.

[16]For example, the standard deviations of the performance test scores for radiation protection, technique and quality control, and positioning for the group tested at work were only about half as large as for the group tested away from work.

Table 8 presents the correlations of the Proficiency Examination with the two types of supervisor's assessment. The correlations in all areas indicate a definite tendency toward agreement between the Proficiency Examination and the supervisor's assessment based on specific situations. The same relationships appear, but more weakly, for the supervisor's assessment based on overall judgments. Again, the validity of the written proficiency examination is supported.

Table 8

Correlations of Proficiency Examination with supervisor's assessments, for those technologists for whom assessment data were available (n = 128).

	Assessment based on specific situations	Assessment based on overall judgments
Total	.45*	.30*
Radiation protection	.23*	.15
Patient care	.29*	.10
Technique and quality control	.39*	.29*
Positioning	.44*	.27*

*Different from zero at the .01 level of statistical significance.

One possible explanation for the higher correlations based on specific situations is that the overall judgments are more influenced by differences between supervisors in their personal and subjective standards for judgment. This explanation would also account for the fact that the assessment based on specific situations is more highly correlated with the performance test than is the assessment based on overall judgments, as can be seen in Table 9. Again, this relationship holds in all four skill areas.

Table 9

Intercorrelation of criterion measures, for all technologists for whom supervisor's assessment data were available (n = 128).

	Performance test with assessment based on specific situations	Performance test with assessment based on overall judgments	Assessment based on specific situations with assessment based on overall judgments
Total	.31*	.11	.53*
Radiation protection	.32*	.26*	.53*
Patient care	.10	-.01	.18
Technique and quality control	.28*	.04	.35*
Positioning	.25*	.06	.54*

*Different from zero at the .01 level of statistical significance.

Validity for credentialed and non-credentialed technologists. Tables 10 through 14 present the results of separate data analyses for the credentialed and non-credentialed technologists. Table 10 shows that, with one or two exceptions, the relationships between the Proficiency Examination and the performance test scores were similar for technologists tested at work and away, in both the non-credentialed and credentialed groups. However, the relationships in the credentialed group are lower than those in the non-credentialed group. A word of caution: the correlations in the first and fourth columns of Table 10 are based on small groups of persons and thus might be quite different in another group of the same number of technologists, chosen in the same manner as those in this study.

Table 10

Correlations of Proficiency Examination with performance test, for credentialed and non-credentialed technologists tested at and away from their workplaces.

	Non-credentialed		Credentialed	
	At Work (n = 14)	Away (n = 50)	At Work (n = 58)	Away (n = 18)
Total	.43	.36	-.03	.09
Radiation protection	.21	.27	.19	.02
Patient care	.44	-.09	.07	.19
Technique and quality control	.22	.44	-.16	-.29
Positioning	.57	.64	.21	-.11

Table 11 presents the mean (average) and standard deviation (spread) of scores on the performance test and the Proficiency Examination, for the four groups of technologists. The standard deviations (shown in parentheses) suggest that differences between groups in the size of the correlations presented in Table 10 may be at least partly the result of differences in the amount of variation within the groups. (If there is little variation in proficiency among the members of a group, the correlation between any two measures of proficiency in that group is likely to be small.)

The mean scores in Table 11 indicate that taking the performance test at an unfamiliar institution tended to lower performance scores for both credentialed and non-credentialed technologists, especially in the areas of patient care and technique and quality control. The technologists tested at work tend to score higher on the performance test than those tested away from their workplaces but there was no such difference on the Proficiency Examination. The advantage of taking the performance test in a familiar setting seems to be smallest in the area of positioning, possibly reflecting a greater standardization of procedures in this area than in the other skill areas.

Table 11

Means and standard deviations (shown in parentheses) on performance test and on Proficiency Examination, for credentialed and non-credentialed technologists taking the performance test at and away from their place of employment.

	Maximum Possible Score	Non-credentialed		Credentialed	
		At Work (n = 14)	Away (n = 50)	At Work (n = 58)	Away (n = 18)
Performance test:					
Total	6.00	3.72 (0.98)	3.02 (0.95)	4.26 (0.75)	3.87 (0.78)
Radiation protection	70	44.1 (10.2)	34.4 (14.7)	50.5 (7.4)	50.3 (5.1)
Patient care	129	78.7 (30.6)	52.1 (27.6)	90.6 (25.5)	78.6 (23.9)
Technique and quality control	164	113.9 (17.4)	97.5 (25.3)	123.9 (11.1)	115.8 (17.4)
Positioning	154	112.2 (19.2)	100.8 (28.1)	118.8 (9.8)	120.0 (11.6)
Proficiency Examination:					
Total	230	135.0 (31.4)	130.5 (27.3)	168.1 (24.4)	186.3 (21.1)
Radiation protection	23	12.4 (3.7)	12.0 (3.5)	16.0 (3.7)	19.1 (3.0)
Patient care	22	15.9 (4.2)	16.9 (2.8)	18.6 (2.4)	19.4 (1.6)
Technique and quality control	80	42.1 (12.0)	39.7 (10.4)	53.4 (11.2)	62.0 (7.8)
Positioning	47	31.7 (6.5)	30.3 (7.9)	39.1 (4.7)	41.0 (3.8)

Tables 12 and 13 present the correlations of the Proficiency Examination with the supervisor's assessments, for the non-credentialed and credentialed technologists separately, as well as for the combined group. In almost all cases, the correlations within the separate groups are smaller than those for the combined group. The correlations for technique and quality control and for radiation protection are higher in the credentialed group than in the non-credentialed group, while those for positioning are higher in the non-credentialed group. The correlations for the total score are similar in the two groups.

Table 12

Correlations of Proficiency Examination with supervisor's assessment based on specific situations, for non-credentialed and credentialed technologists separately.

	Non-credentialed (n = 53)	Credentialed (n = 75)	Combined (n = 128)
Total	.35	.30*	.45*
Radiation protection	.07	.23	.23*
Patient care	.34	.13	.29*
Technique and quality control	.14	.32*	.39*
Positioning	.43*	.32*	.44*

*Different from zero at the .01 level of statistical significance.

Table 13

Correlations of Proficiency Examination with supervisor's assessment based on a single overall judgment for each skill area, for non-credentialed and credentialed technologists separately.

	Non-credentialed (n = 53)	Credentialed (n = 75)	Combined (n = 128)
Total	.14	.26	.30*
Radiation protection	-.05	.11	.15
Patient care	.00	.14	.10
Technique and quality control	.12	.28	.29*
Positioning	.23	.17	.27*

*Different from zero at the .01 level of statistical significance.

Table 14 presents the mean and standard deviation of scores on the two types of supervisor's assessments, separately for the non-credentialed and credentialed technologists for whom supervisor's assessment data were available. For the assessment based on specific situations, the scores have been scaled to a possible range of 0 to 100. A technologist who was judged by his or her supervisor to be capable of performing all the tasks listed on the assessment form, without any assistance or supervision, would receive a score of 100 in all areas. For the assessment based on overall judgments, the scores are reported on a four-point scale, by scoring the supervisor's responses as follows:

0 = Unsatisfactory

1 = In need of improvement

2 = Generally good

3 = Consistently good

4 = Nearly perfect

In interpreting the data reported in Table 14, it is important to keep in mind that the upper half of the table is based on situations specifically selected to present unusual difficulties in one of the four areas, while the lower half of the table is based on supervisors' judgments of the technologists' performance of their regular everyday work.

Table 14

Means and standard deviations (shown in parentheses) on supervisor's assessments, for non-credentialed and credentialed technologists for whom supervisor's assessment data were available.

Assessment based on specific situations (scale of 0 to 100)	Non-credentialed (n = 53)	Credentialed (n = 75)
Total	84 (13)	94 (9)
Radiation protection	88 (19)	97 (7)
Patient care	87 (15)	91 (12)
Technique and quality control	76 (20)	91 (9)
Positioning	87 (13)	93 (11)
Assessment based on overall judgments (scale of 0 to 4)		
Total	2.5 (0.7)	2.8 (0.8)
Radiation protection	2.2 (1.1)	2.7 (0.9)
Patient care	2.9 (0.8)	3.1 (0.9)
Technique and quality control	2.3 (0.9)	2.7 (1.0)
Positioning	2.5 (0.9)	2.8 (0.8)

Results of an alternate scoring formula for the performance test.

Tables 15 and 16 present the correlations involving the performance test scores that resulted from an alternate scoring formula for the performance test. The alternate scoring included only those checkpoints receiving a scoring weight of three (i.e., those judged necessary to safeguard the well-being of the patient). The correlations in Table 15 are very similar to those in the third column of Table 7, and those in Table 16 are very similar to those in the first two columns of Table 9. Thus, the use of the alternate scoring system has very little effect on the correlations of the performance test with the other measures of proficiency and also supports the concurrent validity of the Proficiency Examination.

Table 15

Correlations of Proficiency Examination with performance test, using alternate scoring system for performance test.

	Technologists tested at work (n = 72)	Technologists tested away (n = 68)	Combined (n = 140)
Total	.21	.41*	.45*
Radiation protection	.27	.49*	.43*
Patient care	.25	.11	.20
Technique and quality control	.05	.43*	.35*
Positioning	.39*	.52*	.49*

*Different from zero at the .01 level of statistical significance.

Table 16

Correlations of performance test with supervisor's assessments, using alternate scoring system for performance test, for all technologists for whom supervisor's assessment data were available (n = 128).

	Assessment based on specific situations	Assessment based on overall judgments
Total	.29*	.10
Radiation protection	.31*	.25*
Patient care	.09	-.03
Technique and quality control	.25*	.00
Positioning	.23*	.10

*Different from zero at the .01 level of statistical significance.

The evidence gathered in this study strongly suggests that the written Proficiency Examination is a valid measure of job proficiency in diagnostic radiologic technology, especially in an examinee population that includes technologists not credentialed by ARRT.

Usefulness of Separate Skill Area Scores on the Proficiency Examination

To determine whether the skill area scores on the Proficiency Examination provide any additional valid information that is not provided by the total score, it is useful to examine the correlations of these scores with the criterion measures of proficiency in the specific skill areas. Tables 17, 18, and 19 provide the necessary information. The comparisons of interest are between the correlations in the top row of the table (those for the Proficiency Examination total score) and the correlations connected by the diagonal line (those for the Proficiency Examination subscores). These comparisons indicate that, except in the area of positioning, the total Proficiency Examination score is at least as valid as the skill area scores as an indicator of proficiency in the separate skill areas. This relationship holds for all three criterion measures - the performance test and both types of supervisor's assessment. Therefore, the skill area scores on the Proficiency Examination (except possibly in the area of positioning) do not appear to provide any additional useful and valid information that is not also provided by the total Proficiency Examination score.[17]

[17]The intercorrelations and alpha reliabilities of the Proficiency Examination skill area scores are presented in Appendix F, page 70.

Table 17

Correlations of total score and subscores on Proficiency Examination with total score and subscores on performance test, for all participating technologists (n = 140).

	Performance test:				
Proficiency Examination:	Total	Radiation Protection	Patient Care	Technique and Quality Control	Positioning
Total	.42	.53	.33	.46	.49
Radiation protection	.32	.43	.24	.36	.39
Patient care	.26	.26	.19	.22	.25
Technique and quality control	.34	.45	.25	.37	.39
Positioning	.52	.66	.39	.57	.61

Table 18

Correlations of total score and subscores on Proficiency Examination with total score and subscores on supervisor's assessment based on specific situations, for all technologists for whom supervisor's assessment data were available (n = 128).

	Supervisor's assessment:				
Proficiency Examination:	Total	Radiation Protection	Patient Care	Technique and Quality Control	Positioning
Total	.45	.30	.38	.40	.41
Radiation protection	.33	.23	.28	.27	.30
Patient care	.30	.14	.29	.23	.32
Technique and quality control	.43	.29	.34	.39	.38
Positioning	.47	.31	.41	.39	.44

Table 19

Correlations of total score and subscores on Proficiency Examination with total score and individual items of supervisor's assessment based on overall judgments, for all technologists for whom supervisor's assessment data were available (n = 128).

	Supervisor's assessment:				
Proficiency Examination:	Total	Radiation Protection	Patient Care	Technique and Quality Control	Positioning
Total	.30	.23	.19	.30	.27
Radiation protection	.25	.15	.18	.27	.21
Patient care	.21	.11	.10	.23	.27
Technique and quality control	.30	.21	.21	.29	.26
Positioning	.39	.24	.14	.28	.27

Information Concerning Pass/Fail Cutoff Scores

The information collected in the process of assessing the validity of the Proficiency Examination can also be used to investigate the effects of specifying a pass/fail cutoff score at any given level. Two types of effect are of particular interest. The first of these is the effect on the pass rate - the proportion of examinees who pass the examination. The second is the effect on the validity of the resulting decision procedure - the extent to which the examinees who pass the test show greater actual proficiency than the examinees who fail. The validity of the pass/fail decision procedure can be described by the product-moment correlation between the pass/fail decision procedure and the scores on a criterion measure - the performance test or the supervisor's assessment. This type of correlation is known as a "point-biserial correlation".[18]

The pass rates and the point-biserial correlations resulting from various cutoff scores for the Proficiency Examination total score are presented in Table 20.

---

[18]The point-biserial correlation is not the same as the biserial correlation, which would be entirely inappropriate in this analysis.

Table 20

Effects of selected pass/fail cutoff scores: Proficiency Examination total score.

Proficiency Examination pass/fail cutoff	Non-credentialed			Credentialed			Combined		
	Pass rate	Point-biserial correlation		Pass rate	Point-biserial correlation		Pass rate	Point-biserial correlation	
		Perf. test	Supv. asmt.		Perf. test	Supv. asmt.		Perf. test	Supv. asmt.
85%	2%	.06	.06	21%	-.13	.05	12%	.09	.14
75%	6%	.16	.06	55%	-.04	.26	33%	.28*	.31*
65%	27%	.02	.19	82%	-.04	.38*	56%	.32*	.40*
55%	63%	.23	.18	93%	.12	.20	79%	.34*	.28*
45%	84%	.37*	.24	100%	-	-	93%	.38*	.27*

*Difference on criterion measure between technologists passing and failing the Proficiency Examination is greater than zero at the .01 level of statistical significance.

This information should be interpreted with some caution, for two reasons. First, the pass rates reported in these tables refer to the technologists who participated in the validation study. These technologists were not a systematically or randomly chosen representative sample of some well-defined population of radiologic technologists. While the Proficiency Examination scores of the credentialed technologists suggest that this group does represent the population of ARRT-credentialed technologists, there is no evidence to determine the representativeness of the sample of non-credentialed technologists. Therefore, the pass rates in another group of non-credentialed technologists might be different from those reported here. Second, the use of the point-biserial correlation to indicate the effects of a cutoff score is justified only if the group of technologists includes some technologists who should pass the test and some who should not. Since the participants for this study included some technologists who do not perform all of the tasks on the performance test, the use of the point-biserial correlation seems reasonable, at least for the non-credentialed technologists and for the combined group.

The point-biserial correlations reported in Table 20 are based on the total scores from the Proficiency Examination, the performance test, and the supervisor's assessment based on specific situations. The higher the point-biserial correlation, the more different are the passing and failing technologists on the criterion measure. The cutoff scores investigated were those of 85, 75, 65, 55, and 45 percent correct. Table 20 shows the effect of each of these cutoff scores as applied to the Proficiency Examination total score. The point-biserial correlations reported in Table 20 suggest that fairly good separation between passing and failing groups can be achieved with any cutoff score in the range of 45 to 75 percent.

Tables similar to Table 20 have been constructed for the four separate skill areas. Because the skill area scores were not found to provide any valid and useful information that is not also provided by the total Proficiency Examination score, these tables will not be discussed. They are presented in Appendix F, page 70.

Table 21 shows the actual two-way distribution of total scores on the Proficiency Examination and the performance test, for all 140 participating technologists. The grouping of Proficiency Examination scores into intervals corresponds approximately to the selection of cutoff scores in Table 20. The numbers in Table 21 are the actual numbers of examinees with scores in the specified ranges on both tests. For example, the number 14 in the second row of the table is the number of examinees who scored between 172 and 195 on the Proficiency Examination and between 3.5 and 4.5 on the performance test.

Table 21

Two-way distribution of scores on Proficiency Examination and performance test: all participating technologists.

	Performance test total score					
Proficiency Examination total score	0-1.5	1.5-2.5	2.5-3.5	3.5-4.5	4.5-6.0	Total
196-230	0	2	6	5	2	15
172-195	0	1	12	14	4	31
148-171	2	4	15	10	4	35
124-147	2	7	11	11	2	33
100-123	3	3	8	4	0	18
0- 99	3	3	2	0	0	8
Total	10	20	54	44	12	140

Table 22 shows the results of setting pass/fail cutoff scores on both the Proficiency Examination and the performance test. Each pair of cutoff scores (one cutoff score for each test) produces a set of four numbers, representing the number of examinees who pass both tests, the number of examinees who pass only the Proficiency Examination, the number who pass only the performance test, and the number who fail both tests. For example, the second group of numbers in the top row of Table 22 indicates that if the pass/fail cutoff scores had been set at 85 percent for the Proficiency Examination and 75 percent for the performance test, 6 of the 140 examinees would have passed both tests; 11 examinees would have passed only the Proficiency Examination, 34 would have passed only the performance test, and 89 would have failed both tests. This information is presented separately for credentialed and non-credentialed technologists in Appendix F, pages 71-74.

Table 22

Two-way pass/fail classifications of Proficiency Examination and performance test: all participating technologists.

Proficiency Examination		Performance test 85% P	85% F	75% P	75% F	65% P	65% F	55% P	55% F	45% P	45% F
85%	P	1	16	6	11	10	7	13	4	15	2
	F	3	120	34	89	55	68	84	39	103	20
75%	P	3	43	18	28	30	16	38	8	44	2
	F	1	93	22	72	35	59	59	35	74	20
65%	P	4	75	29	50	45	34	61	18	73	6
	F	0	61	11	50	20	41	36	25	45	16
55%	P	4	107	39	72	57	54	82	29	99	12
	F	0	29	1	28	8	21	15	14	19	10
45%	P	4	126	40	90	64	66	95	35	115	15
	F	0	10	0	10	1	9	2	8	3	7

Table 23 presents the same type of information as Table 22, but Table 23 is based on the alternate scoring system for the performance test (i.e., the scoring system based on only those checkpoints judged necessary to safeguard the well-being of the patient). When the alternate scoring is used, none of the 140 technologists scores as high as 85 percent of the total possible score on the performance test, and only seven score as high as 75 percent. This information is presented separately for credentialed and non-credentialed technologists in Appendix F, pages 75-78.

Table 23

Two-way pass/fail classification of Proficiency Examination and performance test, using alternate scoring system for performance test.

Proficiency Examination		Performance test 85% P	85% F	75% P	75% F	65% P	65% F	55% P	55% F	45% P	45% F
85%	P	0	17	1	16	7	10	12	5	15	2
	F	0	123	6	117	33	90	66	57	91	32
75%	P	0	46	4	42	18	28	34	12	43	3
	F	0	94	3	91	22	72	44	50	63	31
65%	P	0	79	6	73	30	49	52	27	68	11
	F	0	61	1	60	10	51	26	35	38	23
55%	P	0	111	7	104	40	71	68	43	90	21
	F	0	29	0	29	0	29	10	19	16	13
45%	P	0	130	7	123	40	90	76	54	104	26
	F	0	10	0	10	0	10	2	8	2	8

A good indication of the effectiveness of using an examination with a given minimum passing score is the difference between pass rates of two groups: a high-performance group and a low-performance group. These two groups can be defined on the basis of the performance test. For example, in Figure 1, page 46, the high-performance group consists of those technologists with performance test scores of at least 75 percent of the total possible. The low-performance group consists of those technologists with performance test scores of less than 75 percent. Figure 1 presents the Proficiency Examination pass rates for these two groups, as a function of the minimum passing score. Since a higher passing score means that fewer people will pass, both curves slope downward. The vertical distance between the two curves indicates the difference between the pass rates of the two groups. For example, in Figure 1 this difference is fairly large for a passing score of 55 percent, but quite small for a passing score of 85 percent.

Figures 2, 3, and 4 present the same type of information as Figure 1, except that the performance test cutoff used to define the high-performance and low-performance groups has been changed to 65 percent of the total possible performance test score for Figure 2, 55 percent for Figure 3, and 45 percent for Figure 4. Figures 1 through 4 are based on the information presented in Table 22. The circled points are those actually computed from the data; the curves were drawn so as to fit each set of points as closely as possible by a smooth, downward-sloping curve. No graph has been drawn for the first column of Table 22 (performance test cutoff at 85 percent) because the high-performance group would have consisted of only four persons.

Figure 1. Pass rates for high-performance and low-performance groups: performance test cutoff = 75%.

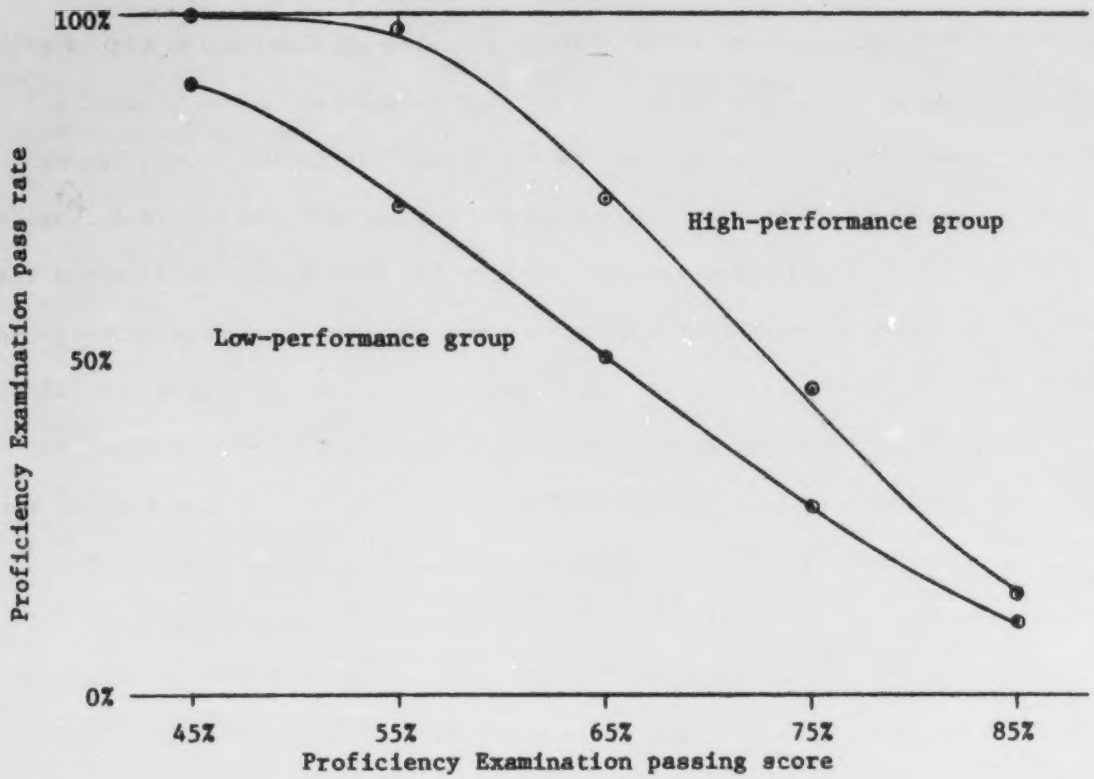


Figure 2. Pass rates for high-performance and low-performance groups: performance test cutoff = 65%.

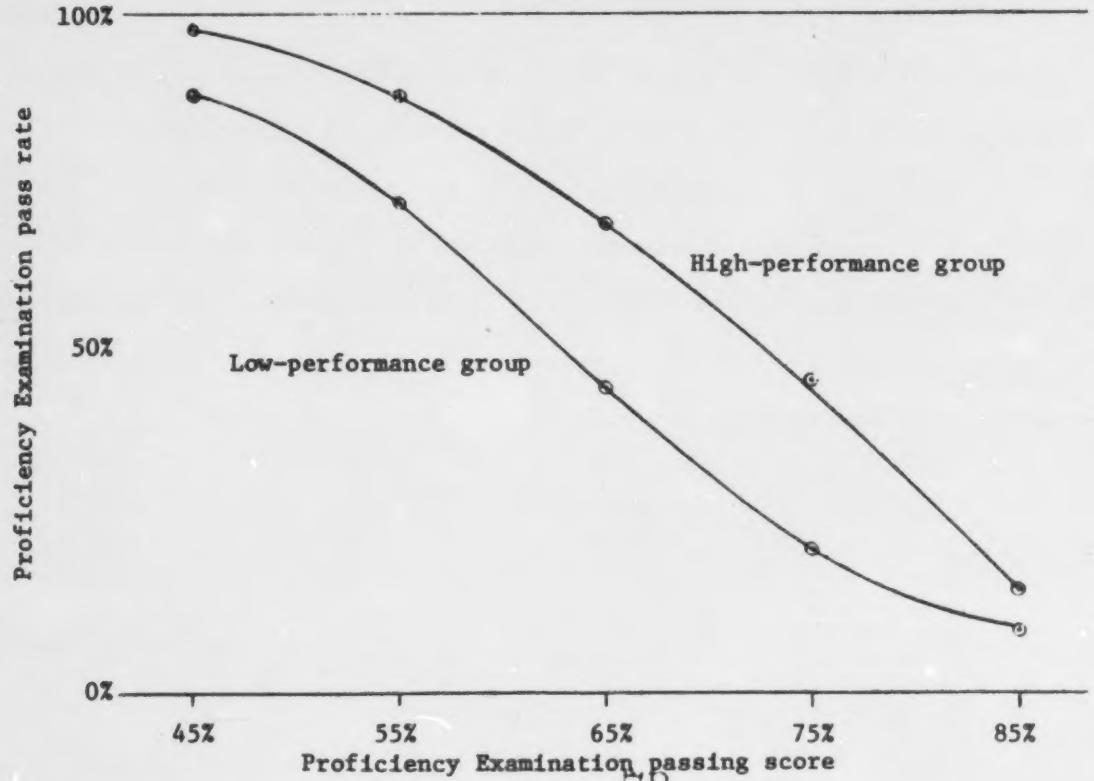


Figure 3. Pass rates for high-performance and low-performance groups: performance test cutoff = 55%.

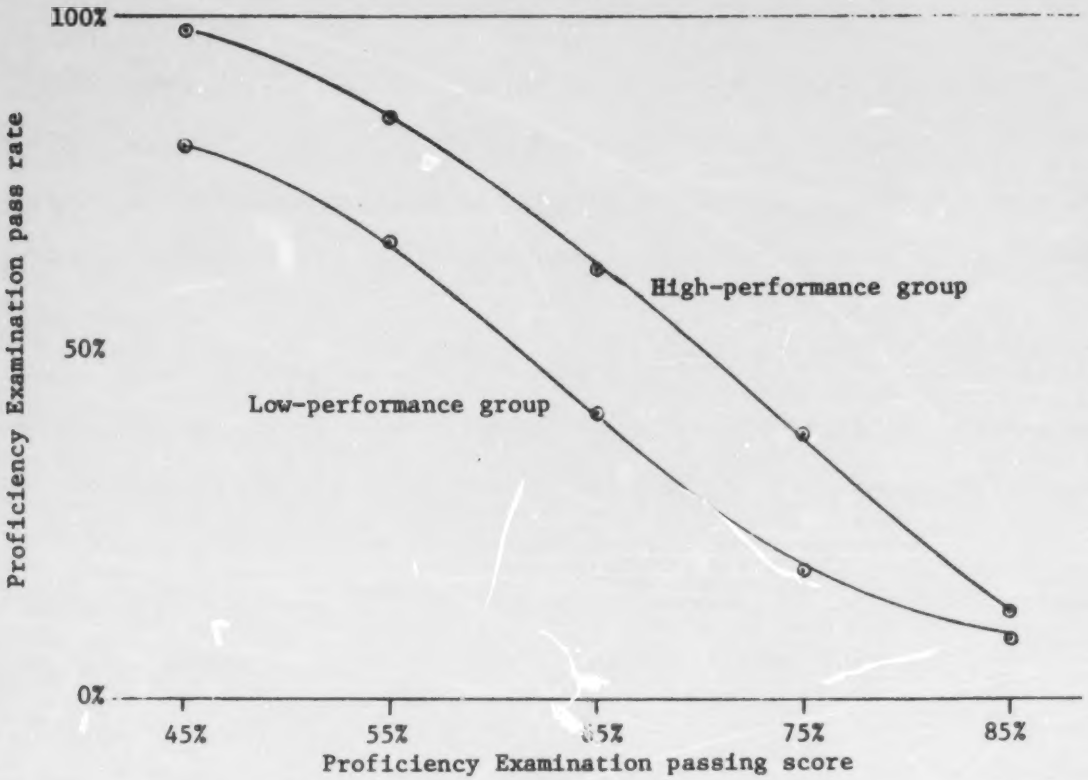
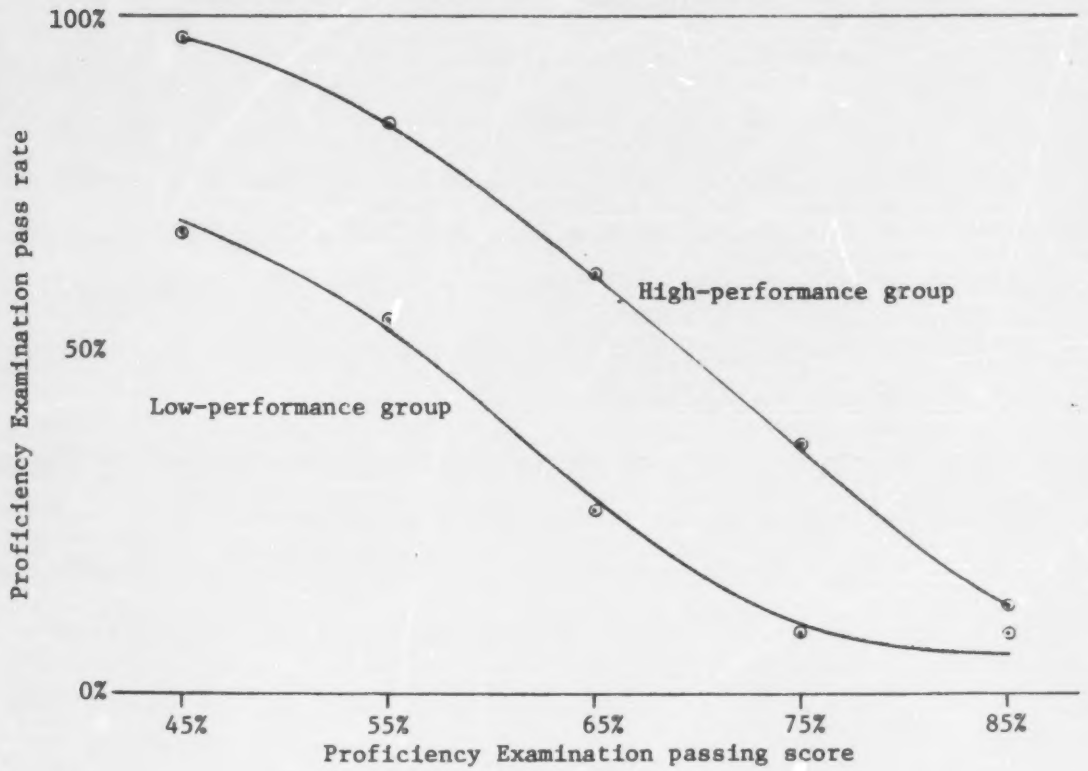


Figure 4. Pass rates for high-performance and low-performance groups: performance test cutoff = 45%.



Figures 1, 2, 3, and 4 thus differ only in the strictness of the definition of "high performance". Figure 1 is based on the strictest definition, which places only 40 of the 140 technologists in the high-performance group. Figure 4 is based on the most lenient definition, which places 118 of the 140 technologists in the high-performance group.

In summary, Figures 1 through 4 show that substantial differences between the pass rates of the high-performance and low-performance groups are achieved when the minimum passing score for the Proficiency Examination is anywhere from 45 to 75 percent.

Any procedure for choosing a pass/fail cutoff score requires some form of judgment as input to the process.[19] The question that must be answered is simply how much proficiency is enough. This question can be answered easily for a single test question (or for a single performance checkpoint, or for a single judgment about the technologist's ability to perform a single specific task). But when scores from several test questions (or performance checkpoints, or specific judgments) are added together, the question is no longer easy to answer. What is needed is a measure by which each technologist's level of proficiency can be classified as acceptable or unacceptable. The supervisor's assessment based on overall judgments can be used to construct such a measure. Although different supervisors may have different standards for acceptability, these differences can be expected to "average out" when the assessments are used to estimate the proportion of technologists whose proficiency is acceptable, using data from a fairly large group of technologists. This proportion can then be compared with the pass rate for the Proficiency Examination, in the same group of technologists, using various pass/fail cutoff scores.

---

[19] Even methods based on decision theory require judgments in the form of utility functions.

Table 24 presents a summary of the supervisors' assessments based on a single overall judgment in each skill area. The numbers represent the number and proportion of technologists who received each rating in each skill area. For example, 14 percent of the technologists were rated as "nearly perfect" in radiation protection, while another 40 percent were rated as "consistently good." Therefore, 54 percent of the technologists for whom supervisors' assessment data were available were rated as "consistently good" or better, in the area of radiation protection. The last row of Table 24 classifies the technologists according to the lowest rating they received in any of the four skill areas. For example, 10 percent of the technologists were rated as "nearly perfect" in all areas, while an additional 30 percent were rated as at least "consistently good" in all areas.

Table 25 shows the results of several pass/fail cutoff scores on the Proficiency Examination in terms of the supervisor's assessment based on overall judgments. This table is similar to Table 22 and can be interpreted in the same way. For example, the block of four numbers in the upper left corner of the table indicates that with a pass/fail cutoff of 85 percent on the Proficiency Examination, 4 technologists who were rated "nearly perfect" in all four skill areas would pass the Proficiency Examination, while 8 would fail; 42 technologists who were rated less than "nearly perfect" in some area would pass the Proficiency Examination, while 74 would fail.

Table 24

Summary of supervisor's assessments based on overall judgments: all technologists for whom supervisor's assessment data were available.*

	"Nearly perfect"	"Consistently good"	"Generally good"	"Needs to improve"	"Unsatisfactory"
Radiation protection	18 (14%)	50 (40%)	45 (36%)	12 (10%)	0 (0%)
Patient care	42 (33%)	48 (38%)	36 (28%)	2 (2%)	0 (0%)
Technique and quality control	22 (17%)	42 (33%)	44 (34%)	20 (16%)	0 (0%)
Positioning	24 (19%)	52 (67%)	43 (34%)	9 (7%)	0 (0%)
Lowest rating	12 (10%)	37 (30%)	50 (40%)	26 (21%)	0 (0%)

*Row totals may vary because of items omitted by supervisors.

Table 25

Two-way pass/fail classification tables of Proficiency Examination and supervisor's assessment based on overall judgments.

Proficiency Examination		Supervisor's Assessment: "Nearly perfect" in all areas		"Consistently good" or better in all areas		"Generally good" or better in all areas	
		Yes	No	Yes	No	Yes	No
85%	P	8	42	25	25	44	6
	F	4	71	24	51	55	20
75%	P	10	74	34	50	70	14
	F	2	39	15	26	29	12
65%	P	11	92	37	66	80	23
	F	1	21	12	10	19	3
55%	P	12	95	39	68	83	24
	F	0	18	10	8	16	2
45%	P	12	105	45	70	92	25
	F	0	8	4	6	7	1

SUMMARY AND CONCLUSIONS

The validity of the Proficiency Examination was investigated by means of a research study in which 140 radiologic technologists took both the Proficiency Examination and a performance test. As an additional criterion measure of job proficiency, supervisors' assessments were obtained for 128 of the technologists. The resulting data indicate that the total score on the Proficiency Examination is a valid indicator of overall job proficiency for radiologic technologists and that separate Proficiency Examination subscores do not provide additional valid information about technologists' proficiency in the specific skill areas of radiation protection, patient care, technique and quality control, and positioning.

The effects of using a pass/fail cutoff at several selected levels were also investigated. The results indicate that any cutoff between 45 and 75 percent of the total possible score on the Proficiency Examination will produce systematic differences in actual performance between the groups passing and failing the Proficiency Examination.

Appendix A. Content specifications for the Proficiency Examination.*

I. Radiation Protection

- A. Monitoring
- B. Shielding
  - 1. Patients
  - 2. Personnel
- C. Units of measure and permissible levels
- D. Radiation effects (e.g., pregnancy)
- E. Technical considerations in reducing radiation exposure and frequency of retakes
  - 1. Filtration
  - 2. Film speeds
  - 3. Types of intensifying screens
  - 4. kVp
  - 5. Collimation

II. Technique and Quality Control

- A. X-ray production
- B. X-ray films and screens
- C. Technique conversions
  - 1. kV
  - 2. mA
  - 3. Distance
  - 4. Grids
  - 5. Time

*From Delineation of Roles and Functions of Diagnostic Radiologic Technology Personnel and Development of Proficiency Tests, pp. 49-52.

6. Filters (copper, aluminum, etc.)

7. Technique adjustments between single and 3-phase

D. Film processing

E. Artifacts

F. Image quality

G. Control of secondary radiation

III. Patient Care

A. Nursing procedures

1. Trauma patient

2. Infectious control

3. Handling of patients

B. Aseptic techniques

C. Emergency procedures

D. Contrast agents (media)

1. Effects and reactions

IV. Positioning

A. Anatomy

1. Normal

2. Pathological

B. Standard projections

1. Magnification

V. Ethics

A. Maintenance of records

B. Confidentiality of communications

C. Professional attitudes

D. Legal responsibilities

E. Interprofessional relationships

VI. Radiograph Equipment and Operator Maintenance

- A. X-ray tubes
  - 1. Tube rating charts
- B. Grids
- C. Standardization of equipment
  - 1. Rectification
  - 2. Spin tops
  - 3. Step wedges
- D. Generators (portable equipment)
- E. Preventive maintenance
- F. Basic electricity
  - 1. Safety and hazards
- G. Automatic processing and maintenance

VII. Special Techniques

- A. Stereo
- B. Body section radiography
- C. Pelvimetry
- D. Image intensification of phototiming
- E. Mobile units

Appendix B. Letter from ETS to radiology departments, and accompanying materials.

EDUCATIONAL TESTING SERVICE PRINCETON, N.J. 08540

Area Code 609
921 - 9000
CABLE-EDUCTESTSVC

•

Center for
Occupational and Professional
Assessment

We would like to request your assistance in a research study being conducted by Educational Testing Service (ETS) to determine the validity of the written proficiency examination for diagnostic radiologic technologist. This examination was developed by ETS under contract with the Bureau of Health Resources Development of HEW and, if accepted by the profession, may be used as an alternate credential for radiologic technologists who lack the formal education necessary for ARRT certification.

The present research is being done under contract with the Division of Associated Health Professions of HEW with the involvement of the American College of Radiology, American Registry of Clinical Radiography Technologists, American Registry of Radiologic Technologists, American Society of Radiologic Technologists and the Bureau of Radiological Health.

The basic purpose of this study is to find out if radiologic technologists who score high on a written proficiency test also do well on a performance or practical test and vice versa. If a positive correlation is found between how technologists answer the written proficiency test questions and how they actually perform on the job, then the written test is likely to be a good indicator of job proficiency.

To conduct this research will require the cooperation of radiologic technologists who will serve as paid volunteer subjects. Each subject will take the written proficiency examination (4 hours long) and a performance test (3 hours long). The performance test involves the actual use of an X-ray machine by the technologist and an actor who will play the role of a patient. Two experienced observers from your staff will be needed to rate each examinee on the performance test. These observers will be paid. We will also request each technologist's supervisor to fill out a confidential rating form to judge the technologist's performance in several different areas of the job.

When all technologists have taken both the performance test and the written examination and all rating forms have been completed by their supervisors, researchers at Educational Testing Service will analyze the data and report the relationships between scores on the written examination, the performance test, and the supervisor's rating form, in each of several areas of job performance. Only these general relationships will be reported publicly; the test scores of an individual technologist will be reported only to him/her, and the supervisors' ratings will not be reported to anybody.

We believe that both the written and performance examinations that will be administered will be of significant educational value to the participating technologists. We therefore hope that your response to this request will be affirmative.

Could you complete the enclosed questionnaire and return it to us by June 25. Should you have any questions please call me collect at 609-921-9000, extension 3508 or Dr. Samuel Livingston at extension 3530.

Thank you very much for your assistance.

Very truly yours,

Amiel Sharon

Amiel T. Sharon, Ph.D.
Program Director

ATS:lil

Enclosures

VALIDATION OF THE PROFICIENCY EXAMINATION FOR

DIAGNOSTIC RADIOLOGIC TECHNOLOGIST

Duties, Responsibilities and Compensation of Participants

A. Test Administrator

The test administrator will be the institutional contact with the project director and will coordinate the collection of all the data within his or her radiology department. This person will:

1. select the two observers for the performance test (possibly acting as one of the observers him or herself);
2. recruit 2 persons to take the role of the patient in the performance test;
3. arrange for the radiographic room to be available for the testing and for the necessary supplies where they are needed;
4. schedule the radiologic technologists for the written and performance tests;
5. arrange for the supervisor of the technologists (examinees) to complete a rating form for each technologist;
6. receive and return all testing materials to Educational Testing Service.

Compensation of the test administrator will be $100 plus $10 per examinee.

B. Performance Test Observer

Two observers will be needed for the performance test. Both should be experienced radiologic technologists. The test administrator may be one of the observers. The function of the observer will be to:

1. assemble and organize the necessary materials for the problem or problems that will be observed;
2. brief the person who will take the role of the patient;
3. give oral instructions to the examinee;
4. restore the radiographic room to specified conditions for each new examinee.

The compensation of the observer will be $10 per hour for time spent in preparation and observation (excluding attendance at the training workshop.) If the test administrator also serves as an observer, her or she will receive compensation for both duties.

C. Examinee

The examinees will be the staff diagnostic radiologic technologists who volunteer to participate in the study. Each examinee will take a performance test (3 hours long) and a written test (4 hours long).

Compensation of each examinee will be $50.

D. Patient

One male and one female actor will be needed to take the roles of the patients in the performance test. The actors may be any adults who are not radiologic technologists, nurses, or physicians. They may be college students majoring in any subject except radiologic technology. The "patient" will be given specific directions regarding how to play the role. No radiographic exposure of the actor will be made on any part of the performance test.

Each patient will be compensated at the rate of $10 per hour.

E. Supervisor

The supervisor(s) of the participating radiologic technologists will be asked to complete a confidential evaluation form for each technologist. No compensation will be provided for this activity.

Following the completion of testing at an institution, the test administrator will complete and sign a payment voucher listing the names, addresses, social security numbers and amounts to be paid to participants in the study. The participants will receive payment approximately three weeks following the receipt of the payment voucher at ETS.

Please return this questionnaire to: Dr. Amiel Sharon
Program Director
Educational Testing Service
Princeton, NJ 08540

Name of Person Completing This Questionnaire _____

Title _____ Telephone _____

Name of Institution _____

Address _____

_____

1. Please indicate how many staff diagnostic radiologic technologists are employed at your institution in each of the categories listed below:

	Number
Credentialed by ARRT	_____
Graduate of an AMA approved program in radiologic technology only	_____
Trained on the job only	_____
Total	_____

2. Approximately how many of the technologists in each of the categories above do you think would be able to participate in the study?

	Number
Credentialed by ARRT	_____
Graduate of an AMA approved program in radiologic technology only	_____
Trained on the job only	_____
Total	_____

3. Would your institution be willing to allow the use of one radiography room for performance testing of technologists? The room will be needed for about five days (not consecutive) during September and October 1976 and will be scheduled at a time most convenient to your institution.

Yes ____ No ____

4. Is there a phantom available at your institution?

Yes ____ No ____

4a. If no, is there a phantom available at a nearby institution?

Yes ____ No ____ Don't know ____

5. Would your institution allow its radiologic technologists to take the written and performance test during or after working hours (the technologists will be paid by ETS for participating in the study)?

Yes ____ No ____

6. Could your institution make available two credentialed expert radiologic technologists (such as the chief technologist) to observe the examinees on the performance test for about five days over a period of two months (the observers will be paid and trained by ETS)?

Yes ____ No ____

7. Would your institution allow ETS personnel to administer the written proficiency examination to its technologists at mutually convenient time(s)?

Yes ____ No ____

8. Please write any comments or questions that you have about this research study below:

PLEASE RETURN THIS QUESTIONNAIRE BY JUNE 25. THANK YOU.

Appendix C. List of Workshop Participants and schedule of activities.

Participants:

Albert M. Snopek, Workshop Organizer
Chairman of Radiologic Technology
Middlesex County College
Edison, New Jersey

Nicholas R. Barraco
School Coordinator
V.A. Hospital School of
Radiologic Technology
West Haven, Connecticut

Norman L. Bossie
Essex County College
Beth. Israel Hospital
Newark, New Jersey

R. Gregory Carroll
Veterans Administration Hospital
Birmingham, Alabama

Elizabeth Curtis
Jersey City Medical Center
Jersey City, New Jersey

Frank E. Ellis
Chief Radiologic Technologist
Williamsport Hospital
Williamsport, Pennsylvania

Susan M. Foster
Instructor, Radiologic Technology
Medical College of Georgia
Augusta, Georgia

J. C. Goodner
Veterans Administration Hospital
Birmingham, Alabama

Nancy A. Gurry
Mount Sinai Hospital
Hartford, Connecticut

Paula Hermer
Montefiore Hospital
Pittsburgh, Pennsylvania

Nancy Lavin
Medical College of Georgia
Augusta, Georgia

Lester Lepley
Chief Technologist Radiology
Franklin Square Hospital
Baltimore, Maryland

Georgene B. McKenzie
Director of Education
Jersey City Medical Center
Jersey City, New Jersey

Sister Mary Rose Mitchell
Chairperson Health Technology
Essex County College
Newark, New Jersey

George P. Osmun
Mount Sinai Hospital
Hartford, Connecticut

Clemens P. Pearson
St. Joseph's Hospital and Medical Center
Phoenix, Arizona

Andre D. Rhodes
Medical College of Virginia
Richmond, Virginia

John Santarsiero
V.A. Hospital School of
Radiologic Technology
West Haven, Connecticut

Sister Agnes Therese
Educational Coordinator
St. Jude Hospital and
Rehabilitation Center
Fullerton, California

Charlotte Todesco
Franklin Square Hospital
Baltimore, Maryland

Michele A. Vaudrevil
St. Joseph's Hospital and Medical Center
Phoenix, Arizona

ASRT Representative:

Marilyn Fay
Director of Continuing Education

HEW Representative:

Margaret A. Wilson
Project Officer

ETS Representatives:

Samuel Livingston
Associate Program Research Scientist
Educational Testing Service

Amiel T. Sharon
Program Director
Educational Testing Service

Mary Vahey
Helene Fuld Medical Center
Trenton, New Jersey

VALIDATION OF PROFICIENCY EXAMINATION FOR
RADIOLOGIC TECHNOLOGY

Agenda for Training Workshop
September 10-11, 1976

September 10, 7:00 P.M.		Dinner Thomas Edison Room Ramada Inn 3050 Woodbridge Avenue Edison, New Jersey
September 11		Amphitheater (2nd Floor) Health Technologies Building Middlesex County College Edison, New Jersey
	8:30 A.M.	Objectives of the Study
	8:45	Chest - Videotape and Discussion
	9:45	IVP - Videotape and Discussion
	10:45	Coffee Break
	11:00	Femoral Neck - Live Demonstration and Discussion
	12:00 P.M.	Lunch
	1:00	Upper GI - Videotape and Discussion
	2:00	Maxilla - Live Demonstration
	2:45	Cervical Spine - Live Demonstration
	3:15	Coffee Break
	3:30	Review and Discussion of All Performance Problems
	5:00	Discussion of Administrative Details
	5:30	Conclusion

Appendix D. Missing data estimation procedure.

Missing checkpoints on the performance test were estimated by the following procedure, which was applied separately for each problem:

To estimate the missing score of examinee i on checkpoint j:

Let $\bar{x}_{i.}$ be the mean score for examinee i on all checkpoints in the problem for which observations of that examinee were made.

Let $\bar{x}_{.j}$ be the mean score on checkpoint j for all examinees of whom observations on that checkpoint were made.

Let $\bar{x}_{..}$ be the mean of all observations actually made of all examinees on all checkpoints in that problem.

Estimate the missing observation by

$$\hat{x}_{ij} = \bar{x}_{..} + (\bar{x}_{i.} - \bar{x}_{..}) + (\bar{x}_{.j} - \bar{x}_{..})$$

If the resulting estimate was greater than one or less than zero, it was replaced by one or zero, respectively. Any examinees who did not attempt a problem or were prevented from finishing the problem were excluded from the missing data estimation procedure for that problem.

Missing responses on the supervisor assessment form were estimated by the same procedure, except that $\bar{x}_{i}$ was the mean for the examinee on all items in the same skill area (radiation protection, patient care, etc.) as the missing item. Examinees for whom no supervisor assessment form was received were not included in this procedure (or in any other procedures involving supervisor assessment data). None of the supervisor assessment forms that were received had more than fifty percent (50%) of the items on

any scale left blank; only four of the forms had more than twenty-five percent (25%) of the items on any scale left blank.

Appendix E. Scoring procedure for the performance test.

The scoring procedure for the performance test was as follows:

1. On each checkpoint, the scores given by the two observers were averaged to form a single checkpoint score for the examinee.

2. The missing data estimation procedure described in Appendix D was used to estimate the value of any checkpoint scores that were missing because of equipment problems or observer oversight.

3. Within each problem, four subscores were computed for each examinee. These subscores, representing the four skill areas (radiation protection, patient care, etc.), were weighted totals of the scores on selected checkpoints. Exception: any examinee with a score of zero on a checkpoint designated as critical received a score of zero for the relevant subscore.

4. Skill area scores were formed by summing the above subscores for each skill area over all six problems.

5. Problem scores were formed in the same manner as the subscores. The weight for each checkpoint was the sum of its weights in all four subscores. Again, a score of zero on a critical checkpoint led to a score of zero for the problem.

6. A total score was computed by dividing each problem score by the maximum possible for that problem and summing the result over all six problems. The purpose of this procedure was to weight each problem equally, regardless of the number of checkpoints it contained.

There was one additional exception to the above procedure. On problem 5 (cervical spine), the subscores for radiation protection and for technique and quality control were not computed, because of the biasing effect of the large amount of missing data resulting from the fact that twenty examinees were prevented from completing the problem. The remaining two subscores - patient care and positioning - were computed, because all the examinees with missing data had erred on at least one crucial checkpoint in these skill areas. Therefore, their subscores for patient care and positioning in problem 5 would have been zero no matter what their performance on the other checkpoints might have been.

An alternate scoring procedure was also used to compute a second set of scores for each participating technologist. The alternate scoring system was the same as that described above, except that only those checkpoints receiving a weight of 3 (indicating an action necessary to safeguard the well-being of the patient) were included.

Appendix F. Supplementary tables.

Table F-1

Intercorrelations and alpha reliabilities of Proficiency Examination scores, for all participating technologists.

	Total	Rad. Prot.	Pat. Care	Tech. & qual. ctl.	Pos.
Total	(.96)	.88	.71	.97	.90
Radiation protection		(.79)	.57	.84	.76
Patient care			(.67)	.66	.55
Technique and quality control				(.92)	.81
Positioning					(.88)

Table F-2

Effects of selected pass/fail cutoff scores: Proficiency Examination subtest score for radiation protection.

Proficiency Examination pass/fail cutoff	Non-credentialed			Credentialed			Combined		
	Pass rate	Point-biserial correlation		Pass rate	Point-biserial correlation		Pass rate	Point-biserial correlation	
		Perf. test	Supv. asmt.		Perf. test	Supv. asmt.		Perf. test	Supv. asmt.
85%	2%	-.02	-.10	24%	.11	.06	14%	.21	.08
75%	8%	.03	.05	49%	.19	.23	30%	.30*	.20
65%	28%	.17	.00	75%	.06	.28	54%	.34*	.19
55%	50%	.23	.01	85%	.17	.19	69%	.37*	.13
45%	67%	.23	-.02	89%	.19	.11	79%	.32*	.08

*Difference on criterion measure between technologists passing and failing the Proficiency Examination subtest is greater than zero at the .01 level of statistical significance.

Table F-3

Effects of selected pass/fail cutoff scores: Proficiency Examination subtest score for patient care.

Proficiency Examination pass/fail cutoff	Non-credentialed			Credentialed			Combined		
	Pass rate	Point-biserial correlation		Pass rate	Point-biserial correlation		Pass rate	Point-biserial correlation	
		Perf. test	Supv. asmt.		Perf. test	Supv. asmt.		Perf. test	Supv. asmt.
85%	25%	-.02	.06	68%	.14	.16	49%	.25*	.18
75%	67%	.03	.28	87%	.05	.05	78%	.14	.21
65%	81%	.09	.20	92%	-.01	.01	87%	.03	.18
55%	91%	.06	.28	99%	.09	.06	95%	.15	.22
45%	97%	.05	.24	100%	-	-	99%	.09	.19

*Difference on criterion measure between technologists passing and failing the Proficiency Examination subtest is greater than zero at the .01 level of statistical significance.

Table F-4

Effects of selected pass/fail cutoff scores: Proficiency Examination subtest score for technique and quality control.

Proficiency Examination pass/fail cutoff	Non-credentialed			Credentialed			Combined		
	Pass rate	Point-biserial correlation		Pass rate	Point-biserial correlation		Pass rate	Point-biserial correlation	
		Perf. test	Supv. asmt.		Perf. test	Supv. asmt.		Perf. test	Supv. asmt.
85%	0%	-	-	14%	.22	.12	8%	.01	.16
75%	3%	.02	-.01	38%	.24	.14	22%	.11	.23*
65%	14%	.07	.07	71%	.16	.25	45%	.27*	.29*
55%	34%	.25	.01	86%	.21	.32*	62%	.33*	.31*
45%	70%	.40*	.18	93%	-.10	.40*	83%	.38*	.33*

*Difference on criterion measure between technologists passing and failing the Proficiency Examination subtest is greater than zero at the .01 level of statistical significance.

Table F-5

Effects of selected pass/fail cutoff scores: Proficiency Examination subtest score for positioning.

Proficiency Examination pass/fail cutoff	Non-credentialed			Credentialed			Combined		
	Pass rate	Point-biserial correlation		Pass rate	Point-biserial correlation		Pass rate	Point-biserial correlation	
		Perf. test	Supv. asmt.		Perf. test	Supv. asmt.		Perf. test	Supv. asmt.
85%	9%	.16	.08	58%	-.02	.25	36%	.24*	.27*
75%	33%	.34*	.34	84%	.15	.28	61%	.42*	.38*
65%	56%	.54*	.42*	95%	.13	.24	77%	.56*	.40*
55%	75%	.57*	.26	100%	-	-	89%	.59*	.25*
45%	88%	.47*	.37*	100%	-	-	94%	.49*	.30*

*Difference on criterion measure between technologists passing and failing the Proficiency Examination subtest is greater than zero at the .01 level of statistical significance.

Table F-6

Two-way pass/fail classification of Proficiency Examination and performance test: non-credentialed technologists only.

Proficiency Examination		Performance test 85% P	85% F	75% P	75% F	65% P	65% F	55% P	55% F	45% P	45% F
85%	P	0	1	0	1	0	1	1	0	1	0
	F	0	63	7	56	15	48	33	30	44	19
75%	P	0	4	1	3	1	3	3	1	4	0
	F	0	60	6	54	14	46	31	29	41	19
65%	P	0	17	2	15	4	13	10	7	14	3
	F	0	47	5	42	11	36	24	23	31	16
55%	P	0	40	6	34	9	31	23	17	31	9
	F	0	24	1	23	6	18	11	13	14	10
45%	P	0	54	7	47	14	40	32	22	42	12
	F	0	10	0	10	1	9	2	8	3	7

Table F-7

Two-way pass/fail classification of Proficiency Examination and performance test: credentialed technologists only.

Proficiency Examination		85%		75%		65%		55%		45%	
		P	F	P	F	P	F	P	F	P	F
85%	P	1	15	6	10	10	6	12	4	14	2
	F	3	57	27	33	40	20	51	9	59	1
75%	P	3	39	17	25	29	13	35	7	40	2
	F	1	33	16	18	21	13	28	6	33	1
65%	P	4	58	27	35	41	21	51	11	59	3
	F	0	14	6	8	9	5	12	2	14	0
55%	P	4	67	33	38	48	23	59	12	68	3
	F	0	5	0	5	2	3	4	1	5	0
45%	P	4	72	33	43	50	26	63	13	73	0
	F	0	0	0	0	0	0	0	0	0	0

Table F-8

Two-way pass/fail classification of Proficiency Examination and performance test, using alternate scoring system for performance test: non-credentialed technologists only.

Proficiency Examination		Performance test 85% P	85% F	75% P	75% F	65% P	65% F	55% P	55% F	45% P	45% F
85%	P	0	1	0	1	0	1	0	1	1	0
	F	0	63	0	63	7	56	21	42	35	28
75%	P	0	4	0	4	1	3	1	3	3	1
	F	0	60	0	60	6	54	20	40	33	27
65%	P	0	17	0	17	3	14	6	11	11	6
	F	0	47	0	47	4	43	15	32	25	22
55%	P	0	40	0	40	7	33	15	25	25	15
	F	0	24	0	24	0	24	6	18	11	13
45%	P	0	54	0	54	7	47	19	35	34	20
	F	0	10	0	10	0	10	2	8	2	8

Table F-9

Two-way pass/fail classification of Proficiency Examination and performance test, using alternate scoring system for performance test: credentialed technologists only.

Proficiency Examination		Performance test 85% P	85% F	75% P	75% F	65% P	65% F	55% P	55% F	45% P	45% F
85%	P	0	16	1	15	7	9	12	4	14	2
	F	0	60	6	34	26	34	45	15	56	4
75%	P	0	42	4	38	17	25	33	9	40	2
	F	0	34	3	31	16	18	24	10	30	4
65%	P	0	62	6	56	27	35	46	16	57	5
	F	0	14	1	13	6	8	11	3	13	1
55%	P	0	71	7	64	33	38	53	18	65	6
	F	0	5	0	5	0	5	4	1	5	0
45%	P	0	76	7	69	33	43	57	19	70	6
	F	0	0	0	0	0	0	0	0	0	0